Available online at www.sciencedirect.com
ScienceDirect
Discrete Mathematics 308 (2008) 1962 – 1972
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

On 2-factors with k components

Gábor N. Sárközy[a, b, 1]

[a] *Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609, USA*
[b] *Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest P.O. Box 63, Budapest H-1518, Hungary*

Received 29 January 2006; received in revised form 10 April 2007; accepted 18 April 2007
Available online 6 May 2007

Abstract

In this paper we study the minimum degree condition for a Hamiltonian graph to have a 2-factor with k components. By proving a conjecture of Faudree et al. [A note on 2-factors with two components, Discrete Math. 300 (2005) 218–224] we show the following. There exists a real number $\varepsilon > 0$ such that for every integer $k \geqslant 2$ there exists an integer $n_0 = n_0(k)$ such that every Hamiltonian graph G of order $n \geqslant n_0$ with $\delta(G) \geqslant (\frac{1}{2} - \varepsilon)n$ has a 2-factor with k components.

Keywords: Hamiltonian cycle; 2-Factor; Minimum degree; Regularity Lemma

1. Introduction

1.1. Notations and definitions

For basic graph concepts see the monograph of Bollobás [3]. $V(G)$ and $E(G)$ denote the vertex-set and the edge-set of the graph G. (A, B, E) denotes a bipartite graph $G = (V, E)$, where $V = A \cup B$, and $E \subset A \times B$. For a graph G and a subset U of its vertices, $G|_U$ is the restriction of G to U. If U_1 and U_2 are two disjoint subsets of the vertices, then $G|_{U_1 \times U_2}$ is the restriction of G to the bipartite graph between U_1 and U_2. $N(v)$ is the set of neighbors of $v \in V$. Hence $|N(v)| = \deg(v) = \deg_G(v)$, the degree of v. $\delta(G)$ stands for the minimum, and $\Delta(G)$ for the maximum degree in G. $\nu(G)$ is the size of a maximum matching in G. C_l denotes the cycle of length l. When A, B are subsets of $V(G)$, we denote by $e(A, B)$ the number of edges of G with one endpoint in A and the other in B. In particular, we write $\deg(v, U) = e(\{v\}, U)$ for the number of edges from v to U. For non-empty A and B,

$$d(A, B) = \frac{e(A, B)}{|A||B|}$$

is the *density* of the graph between A and B. In particular, we write $d(A) = d(A, A) = 2|E(G|_A)|/|A|^2$.

Definition 1. The bipartite graph $G = (A, B, E)$ is ε-regular if

$$X \subset A, \; Y \subset B, \; |X| > \varepsilon|A|, \; |Y| > \varepsilon|B| \quad \text{imply } |d(X, Y) - d(A, B)| < \varepsilon,$$

[1] Research supported in part by the National Science Foundation under Grant no. DMS-0456401.
E-mail address: gsarkozy@cs.wpi.edu.

0012-365X/$ - see front matter
doi:10.1016/j.disc.2007.04.049

otherwise it is ε-irregular. Furthermore, it is (ε, d)-regular if it is ε-regular and $d(A, B) \geqslant d$. Finally, (A, B, E) is (ε, d)-super-regular if it is ε-regular and

$$\deg_G(a) > d|B| \ \forall a \in A, \quad \deg_G(b) > d|A| \ \forall b \in B.$$

1.2. On 2-factors with k components

Let G be a graph on $n \geqslant 3$ vertices. A *Hamiltonian cycle* (*path*) of G is a cycle (path) containing every vertex of G. A *Hamiltonian graph* is a graph containing a Hamiltonian cycle. A graph of order $n \geqslant 3$ is said to be *pancyclic* if it has a cycle of length l for every integer l between 3 and n.

In this paper we study the problem how Hamiltonicity affects the minimum degree condition for the existence of a 2-factor (a union of cycles that span the graph) with k components. A classical result of Dirac [5] asserts that if $\delta(G) \geqslant n/2$, then G is Hamiltonian. This result of Dirac has generated an incredible amount of research, it has been generalized and strengthened in numerous ways (see the survey of Gould [7]). A Hamiltonian cycle is a 2-factor with one component. It is a natural question then to consider the minimum degree condition for the existence of a 2-factor with a certain number of components. Brandt et al. [4] proved that the same bound as in Dirac's Theorem actually guarantees the existence of a 2-factor with k components. More precisely, they proved the following.

Theorem 1 (*Brandt et al. [4]*). *Let k be a positive integer. Then every graph G of order $n \geqslant 4k$ with $\delta(G) \geqslant n/2$ has a 2-factor with k components.*

Here we know that the bound $\delta(G) \geqslant n/2$ is sharp (see [6]). However, we also know that in certain situations if the extra condition of Hamiltonicity is assumed, then we can relax the condition on the minimum degree. A good example for this is the minimum degree condition for pancyclicity.

Theorem 2 (*Amar et al. [1]*). *Let G be a Hamiltonian graph of order n. If $\delta(G) \geqslant (2n+1)/5$, then G is pancyclic or bipartite.*

Here we know that without the extra Hamiltonicity assumption the right bound is again the Dirac bound $\delta(G) \geqslant n/2$.

Faudree et al. [6] raised the question what happens in Theorem 1 if we add the Hamiltonicity assumption. They proved for $k = 2$ that indeed here we can also relax the minimum degree condition.

Theorem 3 (*Faudree et al. [6]*). *Let G be a Hamiltonian graph of order $n \geqslant 6$ with $\delta(G) \geqslant \frac{5}{12}n + 2$. Then G has a 2-factor with two components.*

They also conjectured in [6] that we have a similar situation for 2-factors with k components where $k > 2$; namely we can go below the Dirac bound $\delta(G) \geqslant n/2$. In this paper we prove this conjecture.

Theorem 4. *There exists a real number $\varepsilon > 0$ such that for every integer $k \geqslant 2$ there exists an integer $n_0 = n_0(k)$ such that every Hamiltonian graph G of order $n \geqslant n_0$ with $\delta(G) \geqslant (1/2 - \varepsilon)n$ has a 2-factor with k components, where $k-2$ of the components are C_4's.*

Thus the same constant ε works for every k, this is somewhat stronger than the conjecture in [6], where only the existence of a constant ε_k depending on k was conjectured. However, we note here that in the proof we use the *Regularity method* and thus the constant ε we get is very small. The obtained bound on the minimum degree is probably far from best possible; in fact, the "right" bound might not even be linear (see the discussion in [6]). It would be interesting to determine here the right order of magnitude for this bound.

2. The main tools

In the proof the Regularity Lemma [20] plays a central role. Here we will use the following variation of the lemma. For a proof, see [15].

Lemma 5 (*Regularity Lemma—Degree form*). *For every $\varepsilon > 0$ and every integer m_0 there is an $M_0 = M_0(\varepsilon, m_0)$ such that if $G = (V, E)$ is any graph on at least M_0 vertices and $\delta \in [0, 1]$ is any real number, then there is a partition of the vertex-set V into $l + 1$ sets (so-called clusters) $V_0, V_1, \dots, V_l$, and there is a subgraph $G' = (V, E')$ with the following properties*:

- $m_0 \leqslant l \leqslant M_0$,
- $|V_0| \leqslant \varepsilon |V|$,
- *all clusters V_i, $i \geqslant 1$, are of the same size L*,
- $\deg_{G'}(v) > \deg_G(v) - (\delta + \varepsilon)|V|$ *for all* $v \in V$,
- $G'|_{V_i} = \emptyset$ (*V_i are independent in G'*),
- *all pairs $G'|_{V_i \times V_j}$, $1 \leqslant i < j \leqslant l$, are ε-regular, each with a density 0 or exceeding δ.*

This form can easily be obtained by applying the original Regularity Lemma (with a smaller value of ε), adding to the exceptional set V_0 all clusters incident to many irregular pairs, and then deleting all edges between any other clusters where the edges either do not form a regular pair or they do but with a density at most δ.

An application of the Regularity Lemma in graph theory is now often coupled with an application of the Blow-up Lemma (see [10] for the original, [11] for an algorithmic version and [17,18] for two alternative proofs). Here we use a very special case of the Blow-up Lemma. This asserts that if (A, B) is a super-regular pair with $|A| = |B|$ and $x \in A$, $y \in B$, then there is a Hamiltonian path starting with x and ending with y. More precisely.

Lemma 6. *For every $\delta > 0$ there are $\varepsilon_{BL} = \varepsilon_{BL}(\delta)$, $n_{BL} = n_{BL}(\delta) > 0$ such that if $\varepsilon \leqslant \varepsilon_{BL}$ and $n \geqslant n_{BL}$, $G = (A, B)$ is an (ε, δ)-super-regular pair with $|A| = |B| = n$ and $x \in A$, $y \in B$, then there is a Hamiltonian path in G starting with x and ending with y.*

We will also use some well-known properties of regular pairs. They can be found in [15]. The first one basically says that every regular pair contains a "large" super-regular pair.

Lemma 7 (*Komlós and Simonovits [15, Fact 1.3]*). *Let (A, B) be an (ε, δ)-regular pair and B' be a subset of B of size at least $\varepsilon|B|$. Then there are at most $\varepsilon|A|$ vertices $v \in A$ with $|N(v) \cap B'| < (\delta - \varepsilon)|B'|$.*

The next property says that subgraphs of a regular pair are also regular.

Lemma 8 (*Slicing Lemma, Komlós and Simonovits [15, Fact 1.5]*). *Let (A, B) be an (ε, δ)-regular pair, and, for some $\beta > \varepsilon$, let $A' \subset A$, $|A'| \geqslant \beta|A|$, $B' \subset B$, $|B'| \geqslant \beta|B|$. Then (A', B') is an (ε', δ')-regular pair with $\varepsilon' = \max\{\varepsilon/\beta, 2\varepsilon\}$ and $|\delta' - \delta| < \varepsilon$.*

We will also use two simple Pósa-type lemmas on Hamiltonicity. The second one is the bipartite version of the first one.

Lemma 9 (*Pósa [16]*). *Let G be a graph on $n \geqslant 3$ vertices with degrees $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_n$ such that for every $1 \leqslant k < n/2$ we have $d_k > k$. Then G is Hamiltonian.*

Lemma 10 (*Berge [2, Chapter 10, Theorem 15]*). *Let $G = (A, B)$ be a bipartite graph with $|A| = |B| = n \geqslant 2$ with degrees of vertices in A $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_n$ and with degrees of vertices in B $d'_1 \leqslant d'_2 \leqslant \cdots \leqslant d'_n$. Suppose that for every $1 \leqslant j < n/2$ we have $d_j > j$ and that for every $1 \leqslant k < n/2$ we have $d'_k > k$. Then G is Hamiltonian.*

Finally we will use the following simple fact.

Lemma 11 (*Erdős, Pósa, see Bollobás [3]*). *Let G be a graph on n vertices. Then*

$$\nu(G) \geqslant \min\left\{\delta(G), \frac{n-1}{2}\right\}.$$

3. Outline of the proof

In this paper we use the Regularity Lemma–Blow-up Lemma method again (see [8–14,19]). The method is usually applied to find certain spanning subgraphs in dense graphs. Typical examples are spanning trees (Bollobás-conjecture, see [8]), Hamiltonian cycles or powers of Hamiltonian cycles (Pósa–Seymour conjecture, see [12,13]) or H-factors for a fixed graph H (Alon–Yuster conjecture, see [14]).

Let us take a positive integer $k \geqslant 2$. We may assume that $k \geqslant 3$ since otherwise Theorem 3 implies Theorem 4. We will assume throughout the paper that n is sufficiently large. We will use the following main parameters

$$0 < \varepsilon \ll \delta \ll \alpha \ll 1, \tag{1}$$

where δ depends on α, ε depends on δ and α, and $a \ll b$ means that a is sufficiently small compared to b. We will fix these constants in the beginning of the formal proof of Theorem 4 in the next section.

Let us consider a Hamiltonian graph G of order n with

$$\delta(G) \geqslant (\tfrac{1}{2} - \varepsilon)n. \tag{2}$$

We must show that G has a 2-factor with k components, where at least $k-2$ of the components are C_4's.

First in Section 4.1, in the non-extremal part of the proof, we show this assuming that the following extremal condition does not hold for our graph G. We show later in Section 4.2 that Theorem 4 is true in the extremal case as well. Note that the Hamiltonicity assumption will be used only in a special case of the extremal case.

Extremal Condition (EC) with parameter α: There exist (not necessarily disjoint) $A, B \subset V(G)$ such that

- $|A|, |B| \geqslant (\frac{1}{2} - \alpha)n$, and
- $d(A, B) < \alpha$.

In the non-extremal case, when G does not satisfy EC with parameter α, we apply Lemma 5 for G, with ε and δ as in (1). We get a partition of $V(G') = \bigcup_{0 \leqslant i \leqslant l} V_i$. We define the following *reduced graph* G_r: the vertices of G_r are $p_1, \ldots, p_l$, and we have an edge between vertices p_i and p_j if the pair (V_i, V_j) is ε-regular in G' with density exceeding δ. Thus we have a one-to-one correspondence $f : p_i \to V_i$ between the vertices of G_r and the clusters of the partition. This function f allows us to move from G_r to G' (or G). Since in G', $\delta(G') > (\frac{1}{2} - \varepsilon - (\delta + \varepsilon))n = (\frac{1}{2} - \delta - 2\varepsilon)n$, an easy calculation shows that in G_r we have

$$\delta(G_r) \geqslant (\tfrac{1}{2} - 2\delta)l. \tag{3}$$

Indeed, because the neighbors of $u \in V_i$ in G' can only be in V_0 and in the clusters which are neighbors of p_i in G_r, then for a V_i, $1 \leqslant i \leqslant l$ we have

$$\left(\frac{1}{2} - \delta - 2\varepsilon\right) nL \leqslant \sum_{u \in V_i} \deg_{G'}(u) \leqslant \varepsilon nL + \deg_{G_r}(p_i)L^2.$$

From this using $\varepsilon \leqslant \delta/3$ we get inequality (3):

$$\deg_{G_r}(p_i) \geqslant \left(\frac{1}{2} - \delta - 3\varepsilon\right)\frac{n}{L} \geqslant \left(\frac{1}{2} - 2\delta\right) l.$$

Applying Lemma 11 we can find a matching M in G_r of size at least $(\frac{1}{2} - 2\delta)l$. Put $|M| = m$. Let us put the vertices of the clusters not covered by M into the exceptional set V_0. For simplicity V_0 still denotes the resulting set. Then

$$|V_0| \leqslant 4\delta lL + \varepsilon n \leqslant 5\delta n. \tag{4}$$

Denote the ith pair in $f(M)$ by (V_1^i, V_2^i) for $1 \leqslant i \leqslant m$.

The rest of the non-extremal case is organized as follows. First in Section 4.1.1 in (V_1^1, V_2^1) we find $k-1$ vertex disjoint C_4's with a simple greedy strategy. Indeed, by using ε-regularity, for the first C_4 we take a typical edge (v_1, v_2) in (V_1^1, V_2^1) so v_1 has "many" neighbors in V_2^1, and v_2 has "many" neighbors in V_1^1, and then (v_3, v_4) is an edge between these two neighborhoods. Then (v_1, v_2, v_3, v_4) is a C_4. We repeat this procedure for the other C_4's. We remove these

$k-1$ vertex disjoint C_4's from (V_1^1, V_2^1). Now we just have to find a Hamiltonian cycle on the remaining vertices to have our 2-factor with k components (thus in the non-extremal case actually $k-1$ components are C_4's and one component is a $C_{n-4(k-1)}$). First in Section 4.1.2 we find connecting paths P_i of length 3 in G between the consecutive pairs in the matching $f(M)$ (for $i=m$ the next pair is $i=1$). In Section 4.1.3 we will take care of the various exceptional vertices and make some adjustments by extending some of the connecting paths so that the distribution of the remaining vertices inside each pair in $f(M)$ is perfect, i.e. there are the same number of vertices left in both clusters of the pair. Finally applying Lemma 6 we close the Hamiltonian cycle in each edge (V_1^i, V_2^i) and thus giving a 2-factor with k components.

4. The proof of Theorem 4

We start by fixing the constants. Let

$$\alpha = \frac{1}{100^6} \quad \text{and} \quad \delta = \frac{\alpha^2}{10^6}. \tag{5}$$

Applying Lemma 6 with $\delta/16$ gives us $\varepsilon_{\mathrm{BL}}(\delta/16)$ and $n_{\mathrm{BL}}(\delta/16)$. Let

$$\varepsilon = \min\left\{\frac{\varepsilon_{\mathrm{BL}}(\delta/16)}{2}, \frac{\delta^2}{64}\right\}. \tag{6}$$

Applying Lemma 5 with this ε and $m_0 = 1/\varepsilon$ provides $M_0 = M_0(\varepsilon, m_0)$. Finally let

$$n_0 = n_0(k) = \frac{8kM_0^2 n_{\mathrm{BL}}(\delta/16)}{\varepsilon}. \tag{7}$$

4.1. The non-extremal case

Throughout this section we assume that the extremal condition EC with parameter α does not hold for G. We apply the Regularity Lemma (Lemma 5) for G with ε and δ given in (5), (6) and $m_0 = 1/\varepsilon$, define the reduced graph G_r, and find the matching M in G_r as described above in the outline. First we find $k-1$ C_4's in (V_1^1, V_2^1).

4.1.1. Finding $k-1$ C_4's

Following the outline given above, applying Lemma 7 let us choose a *typical* vertex $v_1 \in V_1^1$ for which we have

$$|N(v_1) \cap V_2^1| \geqslant (\delta - \varepsilon)L \tag{8}$$

(apart from at most εL exceptional vertices most vertices in V_1^1 satisfy this by Lemma 7). Similarly, let us choose a vertex $v_2 \in N(v_1) \cap V_2^1$ for which

$$|N(v_2) \cap V_1^1| \geqslant (\delta - \varepsilon)L \tag{9}$$

(again using Lemma 7 and (8) apart from at most εL exceptional vertices most vertices in $N(v_1) \cap V_2^1$ satisfy this). Choosing $X = N(v_1) \cap V_2^1$ and $Y = N(v_2) \cap V_1^1$, by (8), (9) and $\varepsilon \leqslant \delta/3$ we can apply the regularity condition for X and Y, so in particular we have $d(X, Y) \geqslant \delta - \varepsilon$. Then we can take an arbitrary edge (v_3, v_4) such that $v_3 \in Y$ and $v_4 \in X$ with $v_4 \neq v_2$. Then (v_1, v_2, v_3, v_4) is a C_4. Let us remove this C_4 from (V_1^1, V_2^1), and apply *repeatedly* the above procedure in the leftover of (V_1^1, V_2^1) until we have $k-1$ C_4's. Since in the union of these C_4's we have at most $4(k-1) \leqslant 4k$ vertices, clearly the above procedure goes through since we have $4k \leqslant \varepsilon n/2M_0 \leqslant \varepsilon n/2l \leqslant \varepsilon L$ (using $n \geqslant 8kM_0/\varepsilon$). Indeed, in (8) and (9) we will still have at least $(\delta - 2\varepsilon)L$ remaining vertices even after the removal of the C_4's that we have so far. Butthen, using $\varepsilon \leqslant \delta/3$ we can still apply the regularity condition as above.

We remove these $k-1$ vertex disjoint C_4's from (V_1^1, V_2^1) and for simplicity we keep the notation. Now we just have to find a Hamiltonian cycle on the remaining vertices to have our 2-factor with k components.

4.1.2. Connecting paths

First we have to find connecting paths P_i of length 3 (counting edges) in G between the consecutive edges in the matching $f(M)$ (for $i=m$ the next edge is $i=1$). We are going to use the following fact repeatedly.

Fact 12. *If $x, y \in V(G)$, then there are at least δn internally disjoint paths of length 3 in G connecting x and y.*

Indeed, using (2) we may choose $A \subset N_G(x)$ with $|A| = \lfloor(\frac{1}{2}-\varepsilon)n\rfloor \geqslant (\frac{1}{2}-\alpha)n$ and $B \subset N_G(y)$ with $|B| = \lfloor(\frac{1}{2}-\varepsilon)n\rfloor \geqslant (\frac{1}{2}-\alpha)n$. The fact that EC with parameter α does not hold for G implies $d(A,B)\geqslant\alpha$. From this it follows that we have at least $(\alpha/2)|A|$ vertices v in A, for which we have $\deg(v,B)\geqslant(\alpha/2)|B|$. Indeed, otherwise we would have

$$d(A,B)=\frac{e(A,B)}{|A||B|}<\frac{(\alpha/2)|A||B|+(\alpha/2)|A||B|}{|A||B|}=\alpha,$$

a contradiction. Then using $\delta\leqslant\alpha/20$ we can select greedily a matching of size at least

$$\frac{\alpha}{5}|B|=\frac{\alpha}{5}\left\lfloor\left(\frac{1}{2}-\varepsilon\right)n\right\rfloor\geqslant\frac{\alpha}{20}n\geqslant\delta n,$$

such that each edge has one endpoint in A and one endpoint in B, these endpoints are vertex disjoint from $\{x,y\}$, and from this Fact 12 follows. Indeed, take a vertex $v\in A$ with $\deg(v,B)\geqslant(\alpha/2)|B|$, and select one of these edges to B as the first matching edge. Remove this edge, and apply this repeatedly in the leftover, namely take a vertex $v'\in A$ with $\deg(v',B)\geqslant(\alpha/2)|B|$, and select one of the remaining edges to B as a matching edge. As long as the matching that we have so far covers fewer than $(\alpha/2)|B|$ vertices, we can select the next matching edge. We remove the at most 2 matching edges that have a non-empty intersection with $\{x,y\}$, and we get the desired matching of size at least $(\alpha/5)|B|$.

For the first connecting path P_1 between (V_1^1,V_2^1) and (V_1^2,V_2^2), by using Fact 12 we connect a typical vertex u of V_2^1 (more precisely a vertex u with $\deg(u,V_1^1)\geqslant(\delta-\varepsilon)L$, most vertices in V_2^1 satisfy this by Lemma 7) and a typical vertex w of V_1^2 (so $\deg(w,V_2^2)\geqslant(\delta-\varepsilon)L$) with a path of length 3. To construct the second connecting path P_2 between (V_1^2,V_2^2) and (V_1^3,V_2^3) we just connect a typical vertex of V_2^2 and a typical vertex V_1^3 with a path of length 3 that is vertex disjoint from P_1. Continuing in this fashion, finally we connect a typical vertex of V_2^m with a typical vertex of V_1^1 with a path of length 3 that is vertex disjoint from all the other connecting paths. Note that we can always find these connecting paths that are vertex disjoint from the connecting paths constructed so far. Indeed, the total number of vertices in the union of these paths is at most $4l\leqslant 4M_0\leqslant\varepsilon L$ using $L\geqslant n/2M_0\geqslant 4M_0/\varepsilon$ (where in turn we used $n\geqslant 8M_0^2/\varepsilon$). Then we can find endpoints for the next connecting path that are vertex disjoint from the connecting paths constructed so far since from every cluster we used up only at most εL vertices, so most of the typical vertices from a cluster are still available. Furthermore, when applying Fact 12 to connect the endpoints, since $\varepsilon L\leqslant\varepsilon n\leqslant(\delta/2)n$ we still have $(\delta/2)n$ internally disjoint paths of length 3 connecting the endpoints that are vertex disjoint from the connecting paths constructed so far.

We remove the internal vertices on these connecting paths from the clusters, but we keep the endpoints. For simplicity we keep the notation for the resulting clusters. These connecting paths will be parts of the final Hamiltonian cycle (our kth component in the 2-factor). If the number of remaining vertices (in the clusters and in V_0) is odd, then we take another typical vertex w' of V_1^2 and by using Fact 12 again we extend P_1 by a path of length 3 that ends now with w'. Now the number of remaining vertices is even since we removed three additional vertices, thus we may always assume that this is the case.

4.1.3. Adjustments and the handling of the exceptional vertices

We already have an exceptional set V_0 of vertices in G. We add some more vertices to V_0 to achieve super-regularity. From V_1^i (and similarly from V_2^i) we remove all vertices u for which $\deg(u,V_2^i)<(\delta-\varepsilon)L$. ε-Regularity and Lemma 7 guarantee that at most εL such vertices exist in each cluster V_1^i.

Thus using (4) and $\varepsilon\leqslant\delta$, we still have

$$|V_0|\leqslant 5\delta n+\varepsilon n\leqslant 6\delta n. \tag{10}$$

Since we are looking for a Hamiltonian cycle, we have to include the vertices of V_0 on the Hamiltonian cycle as well. We are going to extend some of the connecting paths P_i, so now they are going to contain the vertices of V_0. Let us consider the first vertex (in an arbitrary ordering of the vertices in V_0) v in V_0. We find a pair (V_1^i, V_2^i) such that either

$$\deg(v, V_1^i) \geqslant \delta L, \tag{11}$$

in which case we say that v and V_1^i are *friendly*, or

$$\deg(v, V_2^i) \geqslant \delta L, \tag{12}$$

in which case we say that v and V_2^i are friendly. In case (11) holds we assign v to the cluster V_2^i, and in case (12) holds we assign v to the cluster V_1^i. In case (11) holds we extend P_{i-1} (for $i = 1$, P_m) inside the pair (V_1^i, V_2^i) by a path of length 3, and in case (12) holds we extend P_i inside the pair (V_1^i, V_2^i) by a path of length 3, so that now in both cases the paths end with v. Indeed, in case (11) holds (it is similar for (12)) consider the endpoint w of P_{i-1} in V_1^i. Choosing $X = N(w) \cap V_2^i$ and $Y = N(v) \cap V_1^i$, by (11), the fact that w was typical and $\varepsilon \leqslant \delta/3$ we can apply the regularity condition for X and Y, so in particular we have $d(X, Y) \geqslant \delta - \varepsilon$. Then we can take an arbitrary edge (v_1, v_2) between X and Y and then (w, v_1, v_2, v) gives us the desired extension of P_{i-1}.

To finish the procedure for v, in case (11) holds we add one more vertex v' to P_{i-1} after v such that $(v, v') \in E(G)$ and v' is a typical vertex of V_1^i, so $\deg(v', V_2^i) \geqslant (\delta - \varepsilon)L$. In case (12) holds we add one more vertex v' to P_i before v such that $(v, v') \in E(G)$, v' is a typical vertex of V_2^i. Thus now v is included as an internal vertex on the extended connecting path P_{i-1} or P_i.

After handling v, we repeat the same procedure for the other vertices in V_0. However, we have to pay attention to several technical details. First, of course in repeating this procedure we always consider the remaining vertices in each cluster; the internal vertices on the extended connecting paths are always removed. For simplicity we keep the notation. Note that the number of remaining vertices is always even during the whole process.

Second, we make sure that we never assign too many vertices of V_0 to any cluster, and thus we never use up too many vertices from any cluster in the matching. First we claim that each $v \in V_0$ is friendly with at least $l/4$ clusters in the matching. Indeed, assume for a contradiction that there were only $c < l/4$ friendly clusters for a $v \in V_0$. Then, since v has fewer than δL neighbors in clusters that are not friendly with v, using (10) and $\varepsilon < \delta < \frac{1}{56}$ we have

$$\deg_G(v) < cL + (2m - c)\delta L + |V_0| \leqslant \frac{l}{4}L + \delta l L + 6\delta n \leqslant \left(\frac{1}{4} + 7\delta\right) n \leqslant \frac{3}{8} n < \left(\frac{1}{2} - \varepsilon\right) n,$$

which is a contradiction with (2). We assign the vertices $v \in V_0$ as evenly as possible to the pairs (in the matching) of the friendly clusters. Since each vertex $v \in V_0$ has at least $l/4$ friendly clusters, each cluster gets assigned at most $4|V_0|/l$ vertices $v \in V_0$. However, as this is proportional to δL, this creates an additional problem, namely as we keep removing vertices we might loose the super-regularity property inside the matching edges, in the worst case it would be possible that we used up all the δL neighbors of a vertex in the other set. Note, that we never loose ε-regularity, the Slicing Lemma (Lemma 8 with $\beta = \frac{1}{2}$) implies that as long as we still have at least half of the vertices remaining in both clusters, the remaining pair is still $(2\varepsilon, \delta/2)$-regular.

Therefore, we do the following periodic *super-regularity updating* procedure inside the pairs. After removing $\lfloor (\delta/8)L \rfloor$ vertices from a pair (V_1^i, V_2^i), we do the following. In the pair (V_1^i, V_2^i) (that is still $(2\varepsilon, \delta/2)$-regular) we find all vertices u from V_1^i (and similarly from V_2^i) for which $\deg(u, V_2^i) < (\delta/2 - 2\varepsilon)|V_2^i|$ (where we consider only the remaining vertices). Consider one such vertex u. Similarly as above using ε-regularity we extend the connecting path P_{i-1} or P_i by a path of length 4 inside the pair (using two vertices from both clusters of the pair so we do not change the difference between the sizes of the clusters in the pair; this fact will be important later) so that it now includes u as an internal vertex (here u plays the role of $v \in V_0$ in the above). By iterating this procedure we can eliminate all of these exceptional u vertices. Then between two updates in a pair (V_1^i, V_2^i), for the degrees of vertices $u \in V_1^i$ (and similarly in V_2^i) we always have

$$\deg(u, V_2^i) \geqslant \left(\frac{\delta}{2} - 2\varepsilon\right) |V_2^i| - \frac{\delta}{8} L \geqslant \left(\frac{\delta}{2} - 2\varepsilon\right) \frac{L}{2} - \frac{\delta}{8} L = \left(\frac{\delta}{8} - \varepsilon\right) L \geqslant \frac{\delta}{16} L,$$

and thus we maintain a super-regularity condition. Furthermore, Lemma 7 implies that we find at most $2\varepsilon L$ exceptional vertices in one cluster in one update. Thus during the whole process the total number of vertices that we use up from a cluster with this super-regularity updating procedure is at most $(64\varepsilon/\delta)L \leqslant \delta L$ using $\varepsilon \leqslant \delta^2/64$.

Returning to the V_0-vertices, using (10), each cluster gets assigned at most $4|V_0|/l \leqslant 24\delta n/l \leqslant 25\delta L$ vertices from V_0 during the whole process. Note that to handle an assigned V_0-vertex we have to use at most two additional vertices from both clusters of the pair where the vertex was assigned. Thus, we use up at most $100\delta L$ vertices from each cluster for handling the vertices in V_0 and an additional δL vertices in the super-regularity updating procedure, so altogether we used up at most $101\delta L$ vertices from each cluster.

After we are done with this, in the remainder of each pair (V_1^i, V_2^i) we have $|V_1^i|, |V_2^i| \geqslant (1-101\delta)L(\geqslant L/2)$ (using $\delta \leqslant \frac{1}{202}$) and the pair is still $(2\varepsilon, \delta/16)$-super-regular. At this point we might have a small difference ($\leqslant 101\delta L$) between the number of remaining vertices in V_1^i and in V_2^i in a pair. Therefore, we have to make some adjustments. For this purpose we will need some facts about G_r. First we will show that G_r satisfies similar structural properties as G.

Fact 13. *EC with parameter $\alpha/2$ does not hold for G_r.*

Indeed, otherwise suppose for a contradiction that there are $A, B \subset V(G_r)$ such that $|A|, |B| \geqslant (\frac{1}{2} - \alpha/2)l$ and $d_{G_r}(A, B) < \alpha/2$. We will show that in this case EC with parameter α would hold for G as well, a contradiction. Consider $f(A)$ and $f(B)$. We have $f(A), f(B) \subset V(G)$ with

$$|f(A)|, |f(B)| \geqslant \left(\frac{1}{2} - \frac{\alpha}{2}\right)(1-\varepsilon)n \geqslant \left(\frac{1}{2} - \alpha\right)n,$$

giving the first condition in the definition of EC with parameter α. For the second condition in the definition, for the number of edges in G between $f(A)$ and $f(B)$ we get the following upper bound:

$$|E(G|_{f(A)\times f(B)})| < \frac{\alpha}{2}|f(A)||f(B)| + (\delta+\varepsilon)|f(A)|n \leqslant \frac{\alpha}{2}|f(A)||f(B)| + 6\delta|f(A)||f(B)| < \alpha|f(A)||f(B)|$$

(using $\delta < \alpha/12$). Here the first term comes from the edges in G' between $f(A)$ and $f(B)$ (they must come from G_r-edges), and the second term comes from the edges in $G \backslash G'$ between $f(A)$ and $f(B)$. Thus indeed EC with parameter α would hold for G, a contradiction, proving Fact 13.

The next fact will be similar to Fact 12.

Fact 14. *For any* (*not necessarily distinct*) *$p, q \in V(G_r)$ there are at least* $(\alpha/90)l$ *internally disjoint alternating* (*with respect to edges in M*) *paths* (*cycles if $p = q$*) *of length* 5 *connecting p and q, where the M-edges are the* 2nd *and* 4th *edges along the paths.*

Indeed, consider the sets $N_{G_r}(p) \cap V(M)$ and $N_{G_r}(q) \cap V(M)$. Let us denote by A the pairs (in M) of the clusters in the first set and by B the pairs of the clusters in the second set. From (3) and $m = |M| \geqslant (\frac{1}{2} - 2\delta)l$ we have $|A|, |B| \geqslant (12 - 6\delta)l$. Using $\delta < \alpha/12$ and Fact 13 we know that $d_{G_r}(A, B) \geqslant \alpha/2$. Then, as in the proof of Fact 12 we can select a matching M' of size at least $(\alpha/10)|B| \geqslant (\alpha/30)l$ from A to B. By throwing away some edges from M', we can find a matching M'' of size at least $(\alpha/90)l$ from A to B such that for any edge $e \in M$ we have at most one edge of M'' that is incident to e. Then the statement of Fact 14 follows. Note that we have

$$\frac{\alpha}{90}l \geqslant \frac{\alpha}{90}m_0 = \frac{\alpha}{90\varepsilon},$$

so using (5) and (6) there are quite many paths guaranteed by Fact 14.

With these preparations, let us take a pair (V_1^i, V_2^i) with a difference $\geqslant 2$ (if one such pair exists), say $|V_1^i| \geqslant |V_2^i| + 2$ (only remaining vertices are considered). Using Fact 14 with $p = q = f^{-1}(V_1^i)$ we can find an alternating path in G_r of length 5 starting and ending with $f^{-1}(V_1^i)$. Let us denote this path by

$$f^{-1}(V_1^i), p_1, p_1', p_2, p_2', f^{-1}(V_1^i),$$

where (p_1, p_1') and (p_2, p_2') are edges in matching M (and thus they correspond to super-regular pairs), and the other three edges of the path are edges in G_r (and thus they correspond to regular pairs). We remove a typical vertex u_1 from

V_1^i (a vertex for which $\deg(u_1, f(p_1)) \geqslant (\delta/2 - 2\varepsilon)|f(p_1)|$, most remaining vertices satisfy this in V_1^i) and add it to $f(p_1')$ (and thus we preserve the super-regularity of $(f(p_1), f(p_1'))$). We remove a typical vertex u_2 from $f(p_1')$ (a vertex for which $\deg(u_2, f(p_2)) \geqslant (\delta/2 - 2\varepsilon)|f(p_2)|$, most remaining vertices satisfy this in $f(p_1')$) and add it to $f(p_2')$ (and thus we preserve the super-regularity of $(f(p_2), f(p_2'))$). Finally we remove a typical vertex u_3 from $f(p_2')$ (a vertex for which $\deg(u_3, V_1^i) \geqslant (\delta/2 - 2\varepsilon)|V_1^i|$, most remaining vertices satisfy this in $f(p_2')$) and add it to V_2^i (and thus we preserve the super-regularity of (V_1^i, V_2^i)). Furthermore, similarly as above in the super-regularity updating when we add a new vertex to a pair (V_1^j, V_2^j), using ε-regularity we extend the connecting path P_{j-1} or P_j by a path of length 4 inside the pair (using two vertices from both clusters of the pair so we do not change the difference between the sizes of the clusters in the pair) so that it now includes the new vertex as an internal vertex. Thus the overall effect of these changes is that the difference $|V_1^i| - |V_2^i|$ decreases by 2, but the other differences $|V_1^j| - |V_2^j|$ do not change for $1 \leqslant j \leqslant m$, $j \neq i$.

Now we are one step closer to the perfect distribution, and by iterating this procedure we can assure that the difference in every pair is at most 1. However, similarly as above we have to make sure that we never use up too many vertices from each cluster in this part of the procedure. Note that altogether we use up at most $10^4\delta n$ vertices in this part of the procedure. We declare a cluster *forbidden* if we used up αL vertices from that cluster. Then from Fact 14 it follows that we can always find an alternating path that does not contain any forbidden clusters assuming $\delta \leqslant \alpha^2/10^6$. Furthermore, as above we perform periodically the super-regularity update inside each pair.

Thus we may assume that the difference in every pair (V_1^i, V_2^i) is at most 1. We consider only those pairs for which the difference is exactly 1, so in particular the number of remaining vertices in one such a pair is odd. Since we have an even number of vertices left it follows that we have an even number of such pairs. We pair up these pairs arbitrarily. If (V_1^i, V_2^i) and (V_1^j, V_2^j) is one such pair with $|V_1^i| = |V_2^i| + 1$ and $|V_1^j| = |V_2^j| + 1$ (otherwise similar), then similar to the construction above, we find an alternating path in G_r of length 5 between V_1^i and V_1^j, and we move a typical vertex of V_1^i through the intermediate clusters to V_2^j.

Thus we may assume that the distribution is perfect, in every pair (V_1^i, V_2^i) we have the same number of vertices ($\geqslant (1 - 2\alpha)L \geqslant L/2$) left in both clusters and all the pairs are still $(2\varepsilon, \delta/16)$-super-regular. Then all the conditions of Lemma 6 are satisfied as from (6) and (7) we have

$$2\varepsilon \leqslant \varepsilon_{\mathrm{BL}}(\delta/16) \quad \text{and} \quad \frac{L}{2} \geqslant \frac{n}{4l} \geqslant \frac{n}{4M_0} > n_{\mathrm{BL}}(\delta/16).$$

Then Lemma 6 closes the Hamiltonian cycle in every pair.

4.2. The extremal case

First we treat two special cases and then we handle the general extremal case.

Case 1: Assume that we have a partition $V(G) = A_1 \cup A_2$ with $(\frac{1}{2} - \alpha)n \leqslant |A_1| \leqslant n/2$ and $d(A_1, A_2) < \alpha^{1/3}$. Thus roughly speaking, $G|_{A_1}$ and $G|_{A_2}$ are almost complete and the bipartite graph between A_1 and A_2 is sparse.

We define *exceptional* vertices $v \in A_i$, $i \in \{1, 2\}$, as

$$\deg(v, A_{i'}) \geqslant \alpha^{1/6}|A_{i'}|, \quad \{i, i'\} = \{1, 2\}.$$

Note that from the density condition $d(A_1, A_2) < \alpha^{1/3}$, the number of exceptional vertices in A_i is at most $\alpha^{1/6}|A_i|$. We remove the exceptional vertices from each set and add them to the set where they have more neighbors. We still denote the sets by A_1 and A_2. Thus using (2) in $G|_{A_i}$, $i \in \{1, 2\}$, it is certainly true that apart from at most $3\alpha^{1/6}|A_i| = 0.03|A_i|$ (using (5)) exceptional vertices all the degrees are at least $(1 - 3\alpha^{1/6})|A_i| = 0.97|A_i|$, and the degrees of the exceptional vertices are at least $0.45|A_i|$.

First by using Lemma 9 (the degree conditions are clearly satisfied) we find a Hamiltonian cycle in $G|_{A_1}$. Thus we only have to find a 2-factor with $k - 1$ components in $G|_{A_2}$. First with a simple greedy strategy (using the degree conditions) we find $k - 2$ C_4's in $G|_{A_2}$, and we remove these cycles from $G|_{A_2}$. In the leftover the conditions of Lemma 9 are still clearly satisfied (with room to spare), and thus there is a Hamiltonian cycle giving us the kth component in the 2-factor.

Case 2: Assume next that there is a partition $V(G) = A_1 \cup A_2$ with $(\frac{1}{2} - \alpha)n \leqslant |A_1| \leqslant n/2$ and $d(A_1) < \alpha^{1/3}$.

Note that in this case from (2) we also have $d(A_1, A_2) > 1 - 2\alpha^{1/3}$. Thus, roughly speaking in this case we have very few edges in $G|_{A_1}$, and we have an almost complete bipartite graph between A_1 and A_2.

In this case a vertex $v \in A_i$, $i \in \{1, 2\}$, is called *exceptional* if it is not connected to most of the vertices in the other set, more precisely if we have

$$\deg(v, A_{i'}) \leqslant (1 - 2\alpha^{1/6})|A_{i'}|, \quad \{i, i'\} = \{1, 2\}.$$

Note that since $d(A_1, A_2) > 1 - 2\alpha^{1/3}$, the number of exceptional vertices in A_i is at most $\alpha^{1/6}|A_i|$. We remove the exceptional vertices from each set and add them to A_2 if they have more neighbors in A_1, and add them to A_1 if they have more neighbors in A_2. We still denote the resulting sets by A_1 and A_2. Assume that $|A_1| \leqslant |A_2|$, so $|A_2| - |A_1| = r$, where $0 \leqslant r \leqslant 3\alpha^{1/6}|A_2| = 0.03|A_2|$. In $G|_{A_1 \times A_2}$ we certainly have the following degree conditions. Apart from at most $3\alpha^{1/6}|A_i| = 0.03|A_i|$ exceptional vertices for all vertices $v \in A_i$, $i \in \{1, 2\}$ we have

$$\deg(v, A_{i'}) \geqslant (1 - 4\alpha^{1/6})|A_{i'}| = 0.96|A_{i'}|, \quad \{i, i'\} = \{1, 2\},$$

and for the exceptional vertices $v \in A_i$, $i \in \{1, 2\}$ we have

$$\deg(v, A_{i'}) \geqslant 0.45|A_{i'}|, \quad \{i, i'\} = \{1, 2\}.$$

Thus note that in $G|_{A_1 \times A_2}$ the degrees of the exceptional vertices are much more than the number of these exceptional vertices. These degree conditions imply the following fact (similar to Fact 12).

Fact 15. *If $x, y \in A_2$, then in $G|_{A_1 \times A_2}$ there are at least $0.2|A_2|$ internally disjoint paths of length 4 connecting x and y.*

Our goal is to achieve $r = 0$. Let us take a Hamiltonian cycle C in G (this is the only place in the whole paper where we use the Hamiltonicity assumption). C must contain at least r edges from $G|_{A_2}$; consider r edges $e_1, \ldots, e_r$ in $C|_{A_2}$. These edges form $r' \leqslant r$ paths. By repeatedly using Fact 15 and connecting endpoints of these paths with vertex disjoint paths of length 4 in $G|_{A_1 \times A_2}$, we find a cycle of length at most $5r$ ($\leqslant 15\alpha^{1/6}|A_2| = 0.15|A_2|$) containing all edges $e_1, \ldots, e_r$ where all other edges are from $G|_{A_1 \times A_2}$ (note that Fact 15 implies that we can always find vertex disjoint paths). We remove this cycle from G and for simplicity let us still denote the resulting sets by A_1 and A_2. Now we have $|A_1| = |A_2|$.

Next with a simple greedy strategy (using the degree conditions) we find $(k - 2)$ C_4's in $G|_{A_1 \times A_2}$, and we remove these cycles as well. Finally in the leftover the conditions of Lemma 10 are still clearly satisfied (with room to spare), and thus there is a Hamiltonian cycle giving us the kth component in the 2-factor.

Extremal Case: Assume finally that the extremal condition EC with parameter α holds, so we have $A, B \subset V(G)$, $|A|, |B| \geqslant (\frac{1}{2} - \alpha)n$ and $d(A, B) < \alpha$. We may also assume $|A|, |B| \leqslant n/2$. We have three possibilities.

- $|A \cap B| < \sqrt{\alpha}n$. The statement follows from Case 1. Indeed, let $A_1 = A$, $A_2 = V(G) \backslash A_1$, then clearly $d(A_1, A_2) < \alpha^{1/3}$ holds.
- $\sqrt{\alpha}n \leqslant |A \cap B| < (1 - \sqrt{\alpha})n/2$. This case is not possible under the given conditions. In fact, otherwise from (2) we would get

$$|A \cap B| \left(\frac{1}{2} - \alpha\right) n \leqslant \sum_{u \in A \cap B} \deg_G(u) = \sum_{u \in A \cap B} \deg_G(u, A \cup B) + \sum_{u \in A \cap B} \deg_G(u, V(G) \backslash (A \cup B))$$
$$\leqslant 2\alpha n^2 + |A \cap B|(|A \cap B| + 2\alpha n),$$

or

$$|A \cap B| \left(\frac{n}{2} - |A \cap B| - 3\alpha n\right) \leqslant 2\alpha n^2,$$

a contradiction under the given conditions.
- $|A \cap B| \geqslant (1 - \sqrt{\alpha})n/2$. The statement follows from Case 2 by choosing $A_1 = A$, $A_2 = V(G) \backslash A_1$, and then $d(A_1) < \alpha^{1/3}$.

This finishes the extremal case and the proof of Theorem 4.

Acknowledgment

Thanks to an unknown referee whose useful remarks improved the presentation.

References

[1] D. Amar, E. Flandrin, I. Fournier, A. Germa, Pancyclism in Hamiltonian graphs, Discrete Math. 89 (1991) 111–131.
[2] C. Berge, Graphs and Hypergraphs, North-Holland, New York, 1973.
[3] B. Bollobás, Extremal Graph Theory, Academic Press, London, 1978.
[4] S. Brandt, G. Chen, R.J. Faudree, R.J. Gould, L. Lesniak, Degree conditions for 2-factors, J. Graph Theory 24 (1997) 165–173.
[5] G.A. Dirac, Some theorems on abstract graphs, Proc. London Math. Soc. 2 (1952) 68–81.
[6] R.J. Faudree, R.J. Gould, M.S. Jacobson, L. Lesniak, A. Saito, A note on 2-factors with two components, Discrete Math. 300 (2005) 218–224.
[7] R.J. Gould, Advances on the Hamiltonian problem—a survey, Graphs and Combin. 19 (2003) 7–52.
[8] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of a packing conjecture of Bollobás, Combinatorics Probab. Comput. 4 (1995) 241–255.
[9] J. Komlós, G.N. Sárközy, E. Szemerédi, On the square of a Hamiltonian cycle in dense graphs, Random Struct. Algorithms 9 (1996) 193–211.
[10] J. Komlós, G.N. Sárközy, E. Szemerédi, Blow-up Lemma, Combinatorica 17 (1997) 109–123.
[11] J. Komlós, G.N. Sárközy, E. Szemerédi, An algorithmic version of the Blow-up Lemma, Random Struct. Algorithms 12 (1998) 297–312.
[12] J. Komlós, G.N. Sárközy, E. Szemerédi, On the Pósa–Seymour conjecture, J. Graph Theory 29 (1998) 167–176.
[13] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of the Seymour conjecture for large graphs, Ann. Combin. 2 (1998) 43–60.
[14] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of the Alon–Yuster conjecture, Discrete Math. 235 (2001) 255–269.
[15] J. Komlós, M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, in: D. Miklős, V.T. Sós, T. Szönyi (Eds.), Combinatorics, Paul Erdős is Eighty, Bolyai Society Mathematical Studies, vol. 2, Budapest, 1996, pp. 295–352.
[16] L. Pósa, A theorem concerning Hamilton lines, Publ. Math. Inst. Hung. Acad. Sci. 7 (1962) 225–226.
[17] V. Rödl, A. Ruciński, Perfect matchings in ε-regular graphs and the Blow-up Lemma, Combinatorica 19 (1999) 437–452.
[18] V. Rödl, A. Ruciński, A. Taraz, Hypergraph packing and graph embedding, Combinatorics Probab. Comput. 8 (1999) 363–376.
[19] G.N. Sárközy, S. Selkow, E. Szemerédi, On the number of Hamiltonian cycles in Dirac graphs, Discrete Math. 265 (2003) 237–250.
[20] E. Szemerédi, Regular partitions of graphs, Colloques Internationaux C.N.R.S. No. 260 - Problèmes Combinatoires et Théorie des Graphes, Orsay, 1976, pp. 399–401.